

## S-7810 WASHINGTON DIGITAL NEWSPAPERS (WDN) PROGRAM

### SOLE SOURCE JUSTIFICATION:

#### Specific Problem or Need

- What is the business need or problem that requires this contract?
  - The Washington Digital Newspapers (WDN) program provides free public access to online historic newspapers published in Washington State and Territory in support of state RCW 27.04.045(8).
  - The current WDN collection contains over 770,000 newspaper pages and will soon reach over 1 million assets.
  - The program uses Library of Congress standards developed for the National Digital Newspaper Program (NDNP) and must meet specific technical guidelines for image file formats and METS/ALTO metadata standards.
  - Digitized pages need to be uploaded to a system that utilizes the METS/ALTO metadata for full-text, advanced search capabilities, and display the original issue, image page, and text articles as they were first published.
  - The website must be scalable to an enterprise level system able to host up to 2+ million newspaper images with searchable xml files containing page structure and full-text, Optical Character Recognition (OCR) data.

#### Sole Source Criteria

- Is the agency proposing this sole source contract because of unavoidable, critical time delays or issues that prevented the agency from completing this acquisition using a competitive process? If so, please describe. For example, if time constraints are applicable, identify when the agency was on notice of the need for the goods and/or service, the entity that imposed the constraints, explain the authority of that entity to impose them, and provide the timelines within which work must be accomplished.

No

- Is the agency proposing this sole source contract because of a geographic limitation? If the proposed contractor is the only source available in the geographical area, state the basis for this conclusion and the rationale for limiting the size of the geographical area selected.

No

- What are the consequences of not having this sole source filing approved? Describe in detail the impact to the agency and to services it provides if this sole source filing is not approved.

In 2018, OSOS posted RFQQ 18-07, resulting in only three responses. One applicant offered to build a new site from scratch with no prior experience supporting large newspaper collections or the required METS/ALTO metadata standards established by the Library of Congress. The second applicant had experience from building an online newspaper portal for the Wyoming State Library but didn't use the Library of Congress' newspaper preservation and metadata standards. Since 2018, the Wyoming State Library has joined NDNP and now uses the national standards and the Veridian platform. The second applicant no longer offers newspaper digitization and hosting services. The third applicant was DL Consulting, offering Veridian, an acceptable platform with reasonable costs, advanced features specifically for newspaper collections, and meets the national METS/ALTO standards. Based on a recent 2023 review of over 20 state awardees using the required NDNP standards, DL Consulting offer the only third-party hosting platform that demonstrates a mature, stable, website that offers advance search using the needed standards for enterprise level digital newspaper collections.

The WDN program is supported by a single library staff member focused on identifying projects for preservation, digitization workflow, metadata processing, content QA and implementing new Azure cloud storage and services, as per state WATech requirements. The time and training needed to develop and implement a new website and/or restructure metadata for a new system would fall to the one staff member, delaying Azure implementation, production workflow, community outreach and scheduled projects.

Without a well-supported, scalable, third party web platform, such as Veridian, the agency risks losing public access to a high-quality user experience, delays in production, loss of metadata and increased impact on local IT resources. OSOS IT staff have recommended a third-party solution, due to limited resources available in-house. Previous attempts to provide a newspaper website by the agency were unsustainable and did not utilize the national standards. The only other platform options identified as possible candidates were open-source platforms developed by the Library of Congress and the University of Oregon requiring additional in-house development by a dedicated web programmer which is not available to us.

What considerations were given to providing opportunities in this contract for small business, including but not limited to unbundling the goods and/or services acquired.

The work required to produce the outcome cannot be taken apart and done by other organizations. The METS/ALTO support and quality assurance is part of the upload process and the features are part of a standard software template specific to the Veridian platform which can be seen in other digital newspaper collections across the country.

Examples of this web template with interactive map, advanced text search, text correction, and information tab features can be seen in these examples: [Library of Virginia](#), [Colorado Historic Newspapers](#), and the [California Digital Newspaper collections](#). DL Consulting is a small company, less than a dozen people, based in Hamilton, New Zealand.

- Describe the unique features, qualifications, abilities or expertise of the contractor proposed for this sole source contract.

Veridian web hosting services provide a system uniquely tailored to enterprise level historic newspaper collections. Search Engine Optimization, full-text search, and filtering capabilities provide a better user experience for greater discoverability and improved results. The search interface allows a greater variety of search methods including an advanced query syntax, interactive maps, calendars, and unique filtering options to narrow results by subject and category. Search terms are highlighted within the page to quickly show topic relevance. The platform also allows public access through mobile devices using a simple QR code:



Veridian software also supports a unique crowd sourced text correction feature, popular with many site visitors, which cannot be moved to other systems. WDN visitors have the option of registering for an account that lets them add subject tags, correct proper names, dates, and events missed by the OCR software in production. WDN registrants can save their search history and track their text correction rates in a “Top Text Correctors” list. This crowd sourced feature improves text searches for all users and makes it easy for researchers to save their work on large projects. All text corrections made by registered users since the website’s debut in January 2019, over 664,608 lines of corrected text, would be lost.

These advanced search features allow WDN visitors to quickly search our collection of 770,000 pages which is expected to grow to approximately 50-80,000 pages per year. Veridian has proven to be scalable to 5-10 million pages as seen in national collections owned by the Swiss, Singapore, and New Zealand national libraries.

DL Consulting has over twenty years of digital newspaper asset management experience. Customer support is very responsive to data issues each time new batches of content are uploaded, guaranteeing successful transfers, compliance with national standards, and data integrity. The integrity of the hosted data is securely maintained through distributed cloud

storage locations in the U.S. which will ensure that the collection is available to the public through the WDN website. The platform also allows site visitors to access the online collection via mobile apps and use accessibility features for the visually impaired. This is an important feature since our Seattle branch is the Washington Talking Book and Braille Library.

- What kind of market research did the agency conduct to conclude that alternative sources were inappropriate or unavailable?

NDNP is a national effort to build newspaper digitization experience and online historic newspaper collections across 50 states plus U.S. Territories. A recent search of 20+ NDNP state awardee online newspaper programs revealed three primary platforms for large collections: In-house solutions developed by university staff, ChronAm software developed by the Library of Congress, and Veridian. Other digital image management systems, used by many NDNP awardees, did not meet the scalability or flexibility we required to meet users' advanced research needs.

In our initial search for a third-party platform under RFQQ 18-07, only three candidates responded. One candidate wanted to build a site from scratch and two of the three did not understand the importance of the NDNP METS/ALTO standards used by the Library of Congress. A mature and tested platform is essential to maintain current access and services unique to the needs of large online newspaper collections. Only DL Consulting had a ready-to-go platform and demonstrates an understanding of the national standards used to create full-text search results.

- Provide a detailed and compelling description of the costs and risks mitigated by contracting with this contractor (i.e. learning curve, follow-up nature).

Costs associated with this type of project should primarily focus on secure storage, a stable platform, and expert knowledge to support regular batch uploads for a growing collection. The costs associated with developing and migrating to a new platform are approximately the same amount as one year of storage and support from the current vendor. By maintaining the current platform and services the agency will save approximately \$20-30,000 in customization fees.

DL Consulting provides an annual invoice based on a scheduled and agreed upon rate at the beginning of the contract. The invoice is based on regular increments of expected growth to cover storage and technical support. There have been no "surprise fees" and the vendor has been up front about any work beyond the contract that might encounter additional fees so there can be no work done without agreement by both parties.

The pricing schedule increases are based on incremental growth of 100,000 pages. The contract explicitly states that "the annual fee will not increase if the collection page count

does not reach the page limit at the end of the contract year. The “increase of annual fee” will also be reduced to a lower rate once the WSL collection reaches one million pages. Pricing provided by this vendor, and based on our previous RFQQ responses, show that the pricing schedule is competitive within the industry and guarantees a set price for the duration of the contract.

- Is the agency proposing this sole source contract because of special circumstances such as confidential investigations, copyright restrictions, etc.? If so, please describe.

No

### **Reasonableness of Cost**

- Since competition was not used as the means for procurement, how did the agency conclude that the costs, fees, or rates negotiated are fair and reasonable? Please make a comparison with comparable contracts, use the results of a market survey, or employ some other appropriate means calculated to make such a determination.

RFQQ 18-07 solicited only three applicants, one with the offer to build a new site from scratch and no experience in the required standards. Their cost estimate was over \$42,000 for the first year including software license, development, data migrations and one year of hosting. The following year would be over \$30,000 annually for license and hosting. The Veridian platform offered a lower rate in annual costs.

The second applicant had experience in hosting the Wyoming State Library digital newspaper collection and understood the technical search requirements needed but did not support the national METS/ALTO standards. Their offer included costs associated with website development, just under \$29,000, and a \$25,000 annual fee for hosting and support, which was competitive among the applicants, but there was no proposed schedule of annual cost increases for expected growth.

DL Consulting’s offer was in range with the other applicants and provided more detailed information about the costs with the growth of the collection. Their costs for customization and branding of the platform were only \$1,800 more for the first year and \$1000 less than the second applicant.